AD-A145 498   INTERIM SCIENTIFIC REPORT GRANT AFOSR-81-0122 1 JUNE   1/1
               1983 - 31 MAY 1984(U) CALIFORNIA UNIV BERKELEY
               OPERATIONS RESEARCH CENTER   R E BARLOW 20 JUL 84
UNCLASSIFIED   AFOSR-TR-84-0728 AFOSR-81-0122   F/G 5/1   NL
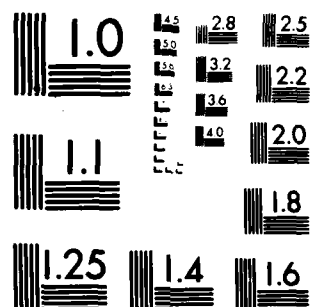
END
DATE
FILMED
10-84
DTIC

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

(4)

AD-A145 498

INTERIM SCIENTIFIC REPORT

GRANT AFOSR-81-0122

1 JUNE 1983 - 31 MAY 1984

20 July 1984
Professor Richard E. Barlow
Operations Research Center
University of California
Berkeley CA 94720

DTIC
SELECTE
SEP 1 0 1934
A

Approved for public release;
distribution unlimited.

84 08 30 039

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED		1b. RESTRICTIVE MARKINGS			
2a. SECURITY CLASSIFICATION AUTHORITY		3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.			
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE					
4. PERFORMING ORGANIZATION REPORT NUMBER(S)		5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 84-0728			
6a. NAME OF PERFORMING ORGANIZATION University of California	6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Air Force Office of Scientific Research			
6c. ADDRESS (City, State, and ZIP Code) Operations Research Center Berkeley CA 94720		7b. ADDRESS (City, State, and ZIP Code) Directorate of Mathematical & Information Sciences, AFOSR, Bolling AFB DC 20332			
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR	8b. OFFICE SYMBOL (If applicable) NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-81-0122			
8c. ADDRESS (City, State, and ZIP Code) Bolling AFB DC 20332		10 SOURCE OF FUNDING NUMBERS			
		PROGRAM ELEMENT NO. 61102F	PROJECT NO. 2304	TASK NO. A5	WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
INTERIM SCIENTIFIC REPORT, GRANT AFOSR-81-0122, 1 JUNE 1983 - 31 MAY 1984.

12. PERSONAL AUTHOR(S)
Richard E. Barlow

13a. TYPE OF REPORT Interim	13b. TIME COVERED FROM 1/6/83 TO 31/5/84	14. DATE OF REPORT (Year, Month, Day) 20 JUL 84	15. PAGE COUNT 9

16. SUPPLEMENTARY NOTATION

17.	COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

The report summarizes research during this period supported by the grant. Topics covered include system reliability, determining sample size for life test experiments, data extractions procedures, and acceptance sampling procedures. Abstracts of papers written during this period are included.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED	
22a. NAME OF RESPONSIBLE INDIVIDUAL CPT Brian W. Woodruff	22b. TELEPHONE (Include Area Code) (202) 767- 5027	22c. OFFICE SYMBOL NM

DD FORM 1473, 84 MAR
83 APR edition may be used until exhausted.
All other editions are obsolete.

84 08 30 039

INTERIM SCIENTIFIC REPORT: AFOSR-81-0122

Research of R. E. Barlow

Perhaps the most significant research, supported in part by the Air Force program, is that on system reliability. Important advances have been made in our understanding of the computational complexity of system reliability problems, especially those problems that can be characterized in terms of graphs or networks. Much of this work, together with new results concerning the more general coherent structure representation, was discussed in "A Survey of Network Reliability." An abstract is attached.

In "Expected Information from a Life Test Experiment," the problem of determining sample size for life test experiments and a methodology based on information expected was given. An abstract is attached.

In "Informative Stopping Rules," the importance and role of information contained in the data extractions procedures was pointed out. Use of this information can result in much better estimators.

Currently, I have written a paper, "A Critique of Deming's Discussion of Acceptance Sampling Procedures," which was presented at the International Conference on Reliability and Quality Control in Columbia, Missouri, 4-8 June 1984. Deming's claim that optimal inspection based on costs should either be *all* or *none* is correct only for certain special prior information. Algorithms for calculating optimal inspection policies have been developed for more general cases. Theorems based on Bayesian posterior analysis partially characterize optimal sampling plans when Deming's all or none statement is incorrect.

Abstracts

[1] A. Agrawal and R. E. Barlow, "A Survey of Network Reliability," ORC 83-5, (July 1983); to appear in forthcoming issue of Operations Research.

We present a brief survey of the current state of the art in network reliability. We survey only exact methods; Monte Carlo methods are not surveyed.
Most network reliability problems are, in the worst case, N-P hard and are, in a sense, more difficult than many standard combinatorial optimization problems.
Although the above sounds very discouraging, there are, in fact, linear and polynomial time algorithms for network reliability problems of special structure.
We review general methods for network reliability computation and discuss the central role played by domination theory in network reliability computational complexity. We also point out the connection with the more general problem of computing the reliability of coherent structures [c.f. Barlow and Proschan, (1981)]. The class of coherent structures contains both directed and undirected networks as well as logic (or fault) trees *without* not gates. This is a rich area for further research.

[2] R. E. Barlow and J. H. Hsiung, "Expected Information from a Life Test Experiment," ORC 82-5, (May 1982); The Statistician, Vol. 32, pp. 35-45, (1983).

Expected information gain as a result of life testing n units for time t is calculated for the time transformed exponential model and a utility function based on entropy. We show that the expected information gain is concave increasing in n and a transform of the test time t . A computer program for calculating expected entropy for the Weibull distribution model is given. This may provide practical guidance in designing life test experiments.

[3] R. E. Barlow and S. W. W. Shor, "Informative Stopping Rules," ORC 84-1, (January 1984); to be presented at the August 1984 ASA Meeting in Philadelphia.

A stopping rule, given data, is informative relative to parameters of interest if it is random and statistically dependent on those parameters. Practical examples, considered in detail, illuminate the role of informative stopping rules and show how they may arise in practice. The discussion is based on the Bayesian approach.

[4] R. E. Barlow and N. D. Singpurwalla, "Assessing the Reliability of Computer Software and Computing Networks: An Opportunity for Partnership with Computer Scientists," ORC 84-7, (July 1984).

Three areas of importance to computer science are highlighted:
(1) Software Reliability; (2) Fault Tolerant Computers; and (3) Network Reliability. This is an up-to-date survey of problems needing further research and results to date.

Research of W. S. Jewell

Research efforts in 1983-4 were concentrated in three areas:

(1) Efficient Calculation of Compound Distributions;

(2) Models of Software Reliability;

(3) Credibility Approximations in Bayesian Prediction.

In 1981, H. Panjer gave a very interesting recursive, exact procedure for calculating the distribution of a random sum of (discrete and positive-valued) random variables, which model occurs in a variety of practical risk problems in engineering and economics. The recursion is based on recursion formulae for underlying counting distribution, which must be Binomial, Poisson, or Negative Binomial (Pascal). Since 1980, we have been examining a variety of models in which this efficient can be used as is, or as a convenient approximation, as in a heterogeneous portfolio, where the number of terms is fixed, the individual "severities" are all different, but there is high probability that an individual severity can be zero. In [1], "Approximating the Distribution of a Dynamic Risk Portfolio," this approach was extended to the situation in which the portfolio has a dynamically changing composition; an alternate interpretation is that each individual risk is, itself, a random sum of random variables. Work in this area continues, with a doctoral thesis investigation of ways in which to incorporate both positive- and negative-valued severities into the Panjer method.

The report, "Bayesian Estimation of Undetected Errors" [2], considers a problem of software reliability in which an unknown number, N, of errors, defects, or bugs exists in a certain product. A number of "inspectors" with unknown competencies inspect the product in parallel, for example, by running independent "Beta" tests. Given the lists of errors found by each inspector, one then wishes to estimate the number of defects *still undetected*. A similar

problem also exists in making a population census of wildlife capture-recapture methods; a maximum-likelihood estimator of N has been known for many years. In this paper we obtain a predictive *distribution* for N , under the assumption that N is Gamma-mixed-Poisson, that errors are equally difficult to detect, and that inspector detection probabilities are independent and Beta-distributed a priori. By casting the problem in a Bayesian framework, we can then give residual-error "guarantees," or examine various inspection decision problems.

The third research report, "Credibility Approximations for Bayesian Prediction of Second Moments" (joint with R. Schnieper) [3], examines the problem of approximating the second moment or variance of the Bayesian predictive distribution. This problem is of practical interest because it is often difficult or extravagant to compute the entire predictive distribution, as when the underlying densities are empirical, and only a few moments are known with certainty. Under the heading of "credibility theory" or "linear-filter theory," linearized approximations of the predictive mean have been known for many years, and have been shown to be exact or robust in many cases of practical interest. Here we extend this work to provide a *simultaneous* estimate of three first and second predictive moments through solutions of a 3x3 linear least-squares system. The results are thus "distribution-free," in that only underlying moments of up to order four are involved. Further, we show that for five often-used likelihoods and priors, the credibility formulas are exact. We believe that they will also be robust approximations in other cases of practical interest.

We have also been planning a research conference for October, 1984, that will be co-sponsored with the Society of Actuaries, on "Credibility Theory and Bayesian Approximation Methods."

Abstracts

[1] W. S. Jewell, "Approximating the Distribution of a Dynamic Risk Portfolio,"
     ORC 83-15, November 1983, Operations Research Center, University of
     California, Berkeley, California.

   In a previous paper, Jewell and Sundt showed how to approximate the
distribution of total losses from a large, fixed, heterogeneous portfolio,
using a recursive algorithm developed by Panjer for the distribution of a
random sum of random variables (a single casualty contract). This paper
extends the approximation procedure to large, dynamic heterogeneous portfolios,
in order to model either a portfolio of correlated casualty contracts, or a
future portfolio, whose composition is not known with certainty.

[2] W. S. Jewell, "Bayesian Estimation of Undetected Errors," ORC 83-11,
     October 1983. To appear, _Proceedings, Second International Meeting
     on Bayesian Statistics_, Valencia, Spain, September 1983.

   An unknown number, N , of errors or defects exist in a certain product,
and I inspectors with unknown competencies are put to work to find the errors.
Given the lists of errors found by each inspector, how can we estimate the
number of undetected errors? A similar problem arises in capture-recapture
sampling in population biology, where the MLE of N , attributed to Petersen,
Chapman, and Darroch, has been known for many years. Our Bayesian model assumes
that N is Gamma-mixed-Poisson, that errors are equally difficult to detect,
and that inspector error detection probabilities are independent and Beta-
distributed, a priori. The predictive density for undetected errors is obtained
as a simple, recursive relationship that gives Negative Binomial tails. The
predictive mode for undetected errors is given by a generalized Petersen-Chapman-
Darroch form involving credibility formulae; as the prior parameter variances
increase without limit, this predictive mode approaches the classical estimator.

[3] W. S. Jewell and R. Schnieper, "Credibility Approximations for Bayesian
     Prediction of Second Moments," ORC 84-3, March 1984; to appear,
     ASTIN Bulletin.

   Credibility theory refers to the use of linear least-squares theory to
approximate the Bayesian forecast of the mean of a future observation; families
are known where the credibility formula is exact Bayesian. Second moment fore-
casts are also of interest, for example, in assessing the precision of the mean
estimate. For some of these same families, the second moment forecast is exact
in linear and quadratic functions of the sample mean. On the other hand, for
the normal distribution with normal-gamma prior on the mean and variance, the
exact forecast of the variance is a linear function of the sample variance and
the squared deviation of the sample mean from the prior mean. Bühlmann has
given a credibility approximation to the variance in terms of the sample mean
and sample variance.
   In this paper we present a unified approach to estimating both first and
second moments of future observations using linear functions of the sample mean
and two sample second moments; the resulting least-squares analysis requires
the solution of a 3x3 linear system, using 11 prior moments from the collective,

and giving joint predictions of all moments of interest. Previously developed special cases follow immediately. For many analytic models of interest, one can replace the 3-dimensional joint prediction with three independent credibility forecasts using the "natural" statistics for each moment.

Research of S. M. Ross

1.   Simulation Analysis of Reliability Systems

Almost all dynamic reliability systems can be modelled as Markov processes either in discrete or continuous time and a basic question is to determine the time, starting from a given initial state, until the process enters a state that is considered failed. As such distributions are usually difficult to evaluate analytically, a simulation analysis was presented by Ross and Schechner in ORC 83-1, "Using Simulation to Estimate First Passage Distributions" (January 1983).

Specifically, they considered a discrete time Markov process $\{X_n,$ $n = 0, 1, \ldots\}$ such that whenever the present state is x the next state is chosen according to the distribution P_x . The initial state i was fixed and for a given set of states A they were interested in estimating the distribution and the mean of N , the number of transitions until the Markov process enters the set A , by use of simulation. By standard techniques such a chain can be simulated until it reaches A --call each such simulation a run. It was then shown that estimators based on

$$N + 1 - \sum_{j=2}^{N+1} P_{X_{j-2}}(A) \; ,$$

where N is the number of steps taken in a given run, and X_j is the jth state in that run and $P_x(A)$ is the probability of going from x to the set A in a single run, has the same mean and smaller variance than the usual estimator N . Hence, the average overall runs of this quantity is a better estimate of $E(N)$ than is the average run size. In addition, a second estimate, based on the observed hazard rate, was given.

Another important problem from a reliability application viewpoint is the estimation of the distribution of the final state. This is important since it represents the failed state and thus repair will depend on it. Such an estimate was presented by working with a modified version of the hazard rate function. Specifically, let $B \subset A$ and define N_B to equal the number of transitions needed to reach B in a run (and thus $N_B = \infty$ if the final state is in $A - B$). Rather than estimating the hazard rate function of N_B, namely $P\{N_B = n \mid N_B \geq n\}$, the modified version $P\{N_B = n \mid N \geq n\}$ was employed, and an estimator based on this was given.

2. Reliability Growth Testing

Consider a complicated system that originally has m defects. Defect i will cause a system failure after a random time that is exponentially distributed with rate λ_i, $i = 1, \ldots, m$. All of the quantities $m, \lambda_1, \ldots, \lambda_m$ are assumed to be unknown. The system is to be run for a time t, with all failures that occur being repaired and the defects that caused the failures being noted. The problem is to estimate the resulting failure rate given that all defects that caused failures in $(0, t)$ are eliminated. Specifically, letting

$$\psi_i(t) = \begin{cases} 1 & \text{if defect } i \text{ does not cause a failure by time } t \\ 0 & \text{otherwise} \end{cases}$$

then we want to estimate

$$\Lambda(t) = \sum_{i=1}^{m} \lambda_i \, \psi_i(t)$$

Under no further assumptions than those listed above, it has been shown (in research that is forthcoming) that $\Lambda(t)$ can be estimated by

$$\sum_{i=1}^{D(t)} \frac{\Lambda_i e^{-\Lambda_i t}}{1-e^{-\Lambda_i t}}$$

where $D(t)$ is the number of observed defects by time t and $\Lambda_1, \ldots, \Lambda_{D(t)}$ are the (determined or estimated) failure rates due to these discovered defects. Stopping times that will enable one to assert with high confidence that $\Lambda(t)$ is within acceptable limits are also being investigated.

The importance of the above is that it deals with an important problem while making a fairly minimal set of assumptions about the specific pattern of growth in reliability.

DATE
ILMED
─8